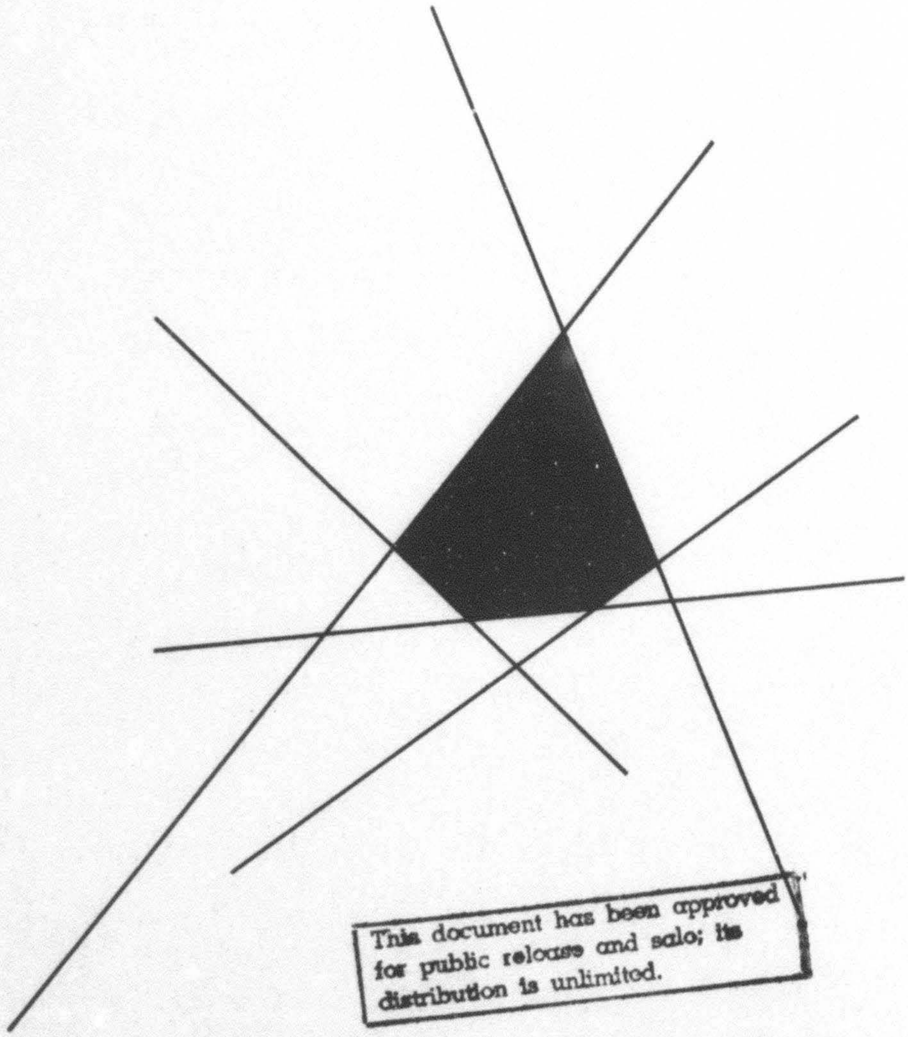
ORC 69-14
JUNE 1969

# CUTTING PLANE METHODS WITHOUT NESTED CONSTRAINT SETS

by
DONALD M. TOPKIS









OPERATIONS RESEARCH CENTER

COLLEGE OF ENGINEERING
UNIVERSITY OF CALIFORNIA · BERKELEY



24

# CUTTING-PLANE METHODS WITHOUT NESTED CONSTRAINT SETS

by

Donald M. Topkis
Department of Industrial Engineering
and Operations Research
University of California, Berkeley

JUNE 1969 ORC 69-14

This research has been partially supported by the Office of Naval Research under Contract Nonr-222(83) and the National Science Foundation under Grant GP-8695 with the University of California.

## ABSTRACT

General conditions are given for the convergence of a class of cutting-plane algorithms without requiring that the constraint sets for the subproblems be sequentially nested. Conditions are given under which inactive constraints may be dropped after each subproblem. Procedures for generating cutting-planes include that of Kelley [4] and a generalization of that used by Zoutendijk [12] and Veinott [9]. For algorithms with nested constraint sets, these conditions reduce to a special case of those of Zangwill [10] for such problems and include as special cases the algorithms of Kelley [4] and Veinott [9]. An arithmetic convergence rate is given.

CUTTING-PLANE METHODS WITHOUT NESTED CONSTRAINT SETS

by

Donald M. Topkis*

I consider the problem of maximizing a real-valued continuous function $f$ over a nonempty closed subset $S$ of $E^n$. It is assumed throughout that one is given a closed subset $T$ of $E^n$ such that $S \subseteq T$ and $f$ attains its maximum over every nonempty closed subset of $T$.[†] The general algorithm to be considered proceeds by setting $T_o = E^n$, and, given $T_k$ as the intersection of $E^n$ and a finite set of closed half-spaces containing $S$, picking $x_k$ to maximize $f$ over $T_k \cap T$, stopping if $x_k \in S$, and otherwise letting $S_k$ be the intersection of $E^n$ and a subset of those half-spaces determining $T_k$ such that $x_k$ maximizes $f$ over $S_k \cap T$, finding a closed half-space $H_k$ containing $S$ but not $x_k$, setting $T_{k+1} = S_k \cap H_k$, and continuing.

$T_k$ will be intersection of $E^n$ and at most $k$ closed half-spaces. It is easily seen that if $T$ is convex and $f$ is pseudo-concave[††] or concave on $T$, then $S_k$ will satisfy the above conditions if $S_k$ is the intersection of all the half-spaces determining $T_k$ for which $x_k$ is on the boundary and any subset

---

* I am grateful to Professor Richard Van Slyke for bringing Reference [6] to my attention and pointing out the relevance of the concept of uniform concavity which allowed the hypotheses of Lemma 7 to be weakened somewhat from an earlier version framed in terms of the matrix of second partial derivatives of $f$ at points of $T$.

† Conditions for this are given in Corollary 11, and it is clearly true when $T$ is compact.

†† A real-valued function $f$ is *pseudo-concave* [7] on a convex set $T \subseteq E^n$ if $f$ is differentiable on $T$ and $(y - x)\cdot\nabla f(x) \le 0$ for $y, x \in T$ implies $f(y) \le f(x)$. A differentiable concave function is pseudo-concave.

of those half-spaces determining $T_k$ for which $x_k$ is an interior point. To be applicable, of course, the subproblems (which have linear constraints when $T$ is a convex polyhedron) must be significantly easier to solve than the original problem, and hence this procedure is essentially confined to problems in which $f$ is either linear, quadratic, or separable in which cases relatively efficient algorithms [1,2,5] exist for the subproblems. The algorithms of Kelley [4] and Veinott [9] set $S_k = T_k$ for all $k$ so each $T_k$ is the intersection of $k$ closed half-spaces and $E^n$. If each $S_k$ is the intersection of $E^n$ and those half-spaces determining $T_k$ for which $x_k$ is a boundary point, then it is seen by Lemma 12 that each $T_k$ is the intersection of no more than $n + 1$ closed half-spaces and $E^n$.

Since $x_k$ maximizes $f$ over $S_k \cap T$, $x_{k+1}$ maximizes $f$ over $T_{k+1} \cap T$, and $S_k \cap T \supseteq T_{k+1} \cap T \supseteq S$,

$$f(x_k) \geq f(x_{k+1}) \geq \max_{x \in S} f(x) . \tag{1}$$

The following slight generalization of an observation of Kelley [4] is clear from (1).

<u>Theorem 1</u>:

If $x_k \in S$ for some $k$, then $x_k$ is optimal. If $x_k \notin S$ for all $k$ and the limit point, $\bar{x}$, of some convergent subsequence of $\{x_k\}$ is feasible, then $\bar{x}$ is optimal.

The conditions of Theorem 10 (which were already indicated to be sufficient for another basic assumption in the first footnote) assure that $\{x_k\}$ is bounded and hence has a convergent subsequence. Thus, the real problem is to find conditions under which the limit point of any convergent subsequence of $\{x_k\}$ is feasible, and this is considered in Section 1.

## 1. GENERAL CONVERGENCE CONDITIONS AND EXAMPLES

A mapping $(a(x),b(x))$ from $T \sim S$ into $E^{n+1}$ with $a(x) \in E^n$ and $b(x) \in E^1$ is a *limiting cutting-plane function* if $S \subseteq H(x) \equiv \{y : a(x)\cdot y \geq b(x)\}$ for all $x \in T \sim S$, $(a(x),b(x))$ is bounded on any bounded subset of $T \sim S$, and for any $\{x_k : k = 1,2, \ldots\} \subseteq T \sim S$ with $\lim_{k\to\infty} x_k = \bar{x} \in T \sim S$ the limit point $(\bar{a},\bar{b})$ of any convergent subsequence of $\{(a(x_k),b(x_k))\}$ satisfies $\bar{a}\cdot\bar{x} < \bar{b}$.[†] The notion of a limiting cutting-plane function was introduced by Zangwill [10]. He applied a generalized version of this notion to a class of cutting-plane algorithms with $S_k = T_k$ to obtain a result for which Corollary 5 is a special case. In addition, he presented Lemma 3 and proved a special case of Lemma 4 with $B = \{0\}$.

Theorem 2:

If $H(x)$ is determined by a limiting cutting-plane function, $H_k = H(x_k)$, $\lim_{i\to\infty} x_{k_i} = \lim_{i\to\infty} x_{j_i} = \bar{x}$, and $x_{j_i} \in H_{k_i}$ for all $i$, then $\bar{x}$ is optimal.

Proof:

Since $(a(x),b(x))$ is a limiting cutting-plane function and $x_{j_i} \in H_{k_i} = H\left(x_{k_i}\right)$ for all $i$,

$$(2) \qquad a\left(x_{k_i}\right)\cdot x_{j_i} \geq b\left(x_{k_i}\right) > a\left(x_{k_i}\right)\cdot x_{k_i} \quad \text{for all } i .$$

If $(\bar{a},\bar{b})$ is the limit point of any convergent subsequence of $\left\{\left(a\left(x_{k_i}\right),b\left(x_{k_i}\right)\right)\right\}$, then it follows from (2) that $\bar{a}\cdot\bar{x} = \bar{b}$. Hence, $\bar{x}$ is feasible since $(a(x),b(x))$

---

[†] If $T$ is convex and a cutting-plane function exists, then it is easily seen that $S$ is convex.

is a limiting cutting-plane function, and so $\bar{x}$ is optimal by Theorem 1. //

The following two lemmas give examples of limiting cutting-plane functions. The function of Lemma 3 was introduced by Kelley [4]. The special case of the function of Lemma 4 with $B = \{0\}$ was introduced by Zoutendijk [12] although he gave no proof for his algorithm, and a slightly modified version of his algorithm was proven to converge by Veinott [9]. Throughout, suppose that $S = \{x : G(x) \geq 0\}$ where $G(x)$ is a real-valued continuous function on $T$.

Lemma 3:

Suppose that there exists a function $\mu(x)$ from $T \sim S$ into $E^n$ which is bounded on any bounded subset of $T \sim S$ and such that $G(y) \leq G(x) + \mu(x)\cdot(y - x)$ for all $x \in T \sim S$ and $y \in S$. Then $(\mu(x), \mu(x)\cdot x - G(x))$ is a limiting cutting-plane function.

Proof:

Clearly, $S \subseteq H(x)$. Pick $\{x_k\} \subseteq T \sim S$ with $\lim_{k\to\infty} x_k = \bar{x} \in T \sim S$. Then the limit point of any convergent subsequence of $\{(\mu(x_k), \mu(x_k)\cdot x_k - G(x_k))\}$ takes the form $(\bar{\mu}, \bar{\mu}\cdot\bar{x} - G(\bar{x}))$. But $G(\bar{x}) < 0$ since $\bar{x} \in T \sim S$ so $\bar{\mu}\cdot\bar{x} < \bar{\mu}\cdot\bar{x} - G(\bar{x})$. //

Kelley [4] observed that if $T$ is convex and $G(x) = \min_{1\leq i\leq m} g_i(x)$ where each $g_i$ is differentiable and concave on $T$ then $\mu(x) = \nabla g_{i(x)}(x)$ satisfies the conditions of Lemma 3 if $i(x)$ is chosen such that $G(x) = g_{i(x)}(x)$.

Lemma 4:

Suppose there exists $t$ with $G(t) > 0$, $S$ is convex, and for $x \in T \sim S$ define $\lambda(x) = \sup\{\lambda : \lambda x + (1 - \lambda)t \in S\}$ and set $w(x) = \alpha(x)x + (1 - \alpha(x))t$ for any $\alpha(x) \in [\lambda(x), 1]$ with $\frac{G(w(x))}{G(x)} \in B$ where $B$ is a nonempty subset of $[0,1]$. Suppose also that for $P = \{w :$ there exists $x \in T \sim S$ with $w(x) = w\}$ there

exists a function $\mu(w)$ from $P$ into $E^n$ which is bounded and bounded away from $0$ on any bounded subset of $P$ and such that $0 \leq G(w) + \mu(w)\cdot(y - w)$ for all $w \in P$ and $y \in S$. Then $\lambda(x) \in (0,1)$ for all $x \in T \sim S$ and $(\mu(w(x)), \mu(w(x))\cdot w(x) - G(w(x)))$ is a limiting cutting-plane function.

Proof:

Pick any $x \in T \sim S$. If $\lambda(x) > 1$ then there exists $z \in S$ and $\gamma \in (0,1)$ with $x = \gamma t + (1 - \gamma)z$ which contradicts the convexity of $S$. Also, $\lambda(x)$ cannot equal 1 since this would imply that $x \in S$ because $S$ is closed, so $\lambda(x) < 1$. Since $G(t) > 0$, it then follows that $\lambda(x) \in (0,1)$.

Now pick $\{x_k\} \subseteq T \sim S$ with $\lim_{k\to\infty} x_k = \bar{x} \in T \sim S$. Then the limit point of any convergent subsequence of $\{(\mu(w(x_k)), \mu(w(x_k))\cdot w(x_k) - G(w(x_k)))\}$ takes the form $(\bar{\mu}, \bar{\mu}\cdot\bar{w} - G(\bar{w}))$ where $\bar{\mu} \neq 0$. Let $\bar{\alpha}$ be the limit point of any convergent subsequence of the corresponding subsequence of $\{\alpha(x_k)\}$. Then $\bar{w} = \bar{\alpha}\bar{x} + (1 - \bar{\alpha})t$, $\bar{\alpha} \in [0,1]$, and $\bar{w} \in \overline{E^n \sim S}$ imply $\bar{\alpha} > 0$ and $G(\bar{w}) \leq 0$. But since $t$ is in the interior of $S$, $\bar{\mu} \neq 0$, and clearly

$$0 \leq G(\bar{w}) + \bar{\mu}\cdot(y - \bar{w}) \quad \text{for all } y \in S,$$

it follows that

$$0 < G(\bar{w}) + \bar{\mu}\cdot(t - \bar{w}) \tag{3}$$

and so by (3) and $G(\bar{w}) \leq 0$

(4) $$0 < \bar{\mu}\cdot(t - \bar{w}) .$$

But $t - \bar{w} = \bar{\alpha}(t - \bar{x})$ and

(5) $$\bar{x} - \bar{w} = (1 - \bar{\alpha})(\bar{x} - t) = -\left(\frac{1 - \bar{\alpha}}{\bar{\alpha}}\right)(t - \bar{w})$$

and so by (4) and (5) if $\bar{\alpha} < 1$

(6) $$0 > \bar{\mu}\cdot(\bar{x} - \bar{w})$$

and by (6) and $G(\bar{w}) \leq 0$

(7) $$0 > G(\bar{w}) + \bar{\mu}\cdot(\bar{x} - \bar{w}) .$$

If $\bar{\alpha} = 1$ then $\bar{w} = \bar{x}$ and (7) still holds.

Since by hypothesis $S \subseteq H(x)$ for all $x \in T \sim S$ , the proof is complete. //

Suppose $G(x) = \min_{1 \leq i \leq m} g_i(x)$ where each $g_i$ is a real-valued differentiable function on $E^n$ . Veinott [9] showed that when $B = \{0\}$ , each $g_i$ is quasi-concave† on $E^n$ , $i(x)$ is any $i$ such that $G(x) = g_{i(x)}(x)$ , and $G(x) = 0$ implies $\nabla g_{i(x)}(x) \neq 0$ , then for $\mu(x) = \nabla g_{i(x)}(x)$ the hypotheses about $\mu(x)$ in Lemma 4 are satisfied. For arbitrary $B \subseteq [0,1]$ it is easily seen that if each $g_i(x)$ is concave and $\mu(x) = \nabla g_{i(x)}(x)$ (with $i(x)$ as above) then this $\mu(x)$ satisfies the hypotheses of Lemma 4. The method of Lemma 4 with $B = (0,1)$ and $\alpha(x) \in (\lambda(x),1)$ would seem to compare favorably with Kelley's method (which requires $\alpha(x) = 1$ , but does not assume the existence of $t$ with $G(t) > 0$) because the cutting-plane for $x$ is generated by a point "closer" to $S$ than $x$ , and the advantage over Veinott's method is that with $B = \{0\}$ finding $\alpha(x)$ *exactly* is generally not

†A real-valued function $f$ is *quasi-concave* [7] on a convex set $T$ if $\{x : f(x) \geq \alpha , x \in T\}$ is convex for all real $\alpha$ . A concave function is quasi-concave.

always possible (although here one would need each $g_i$ to be concave while he only required them to be quasi-concave). Finally, observe that when $B = \{1\}$ and $\alpha(x) = 1$ the method of Lemma 4 is identical with that of Lemma 3.

The following is an immediate consequence of Theorem 2 by letting $j_i = k_{i+1}$.

Corollary 5:

If $H(x)$ is determined by a limiting cutting-plane function, $H_k = H(x_k)$, and $S_k = T_k$, then the limit point of any convergent subsequence of $\{x_k\}$ is optimal.

Corollary 5 together with Lemma 3 proves Kelley's algorithm [4], and Corollary 5 together with Lemma 4 proves Veinott's algorithm [9] and a more general class of cutting-plane algorithms.

The following result paves the way for generating cutting-plane algorithms in which inactive constraints (i.e., half-spaces determining $T_k$ for which $x_k$ is an interior point) may be dropped after each subproblem. It follows immediately from Theorem 2 by letting $j_i = k_i + 1$ since $x_{k_i+1} \in T_{k_i+1} \subseteq H_{k_i}$.

Corollary 6:

If $H(x)$ is determined by a limiting cutting-plane function, $H_k = H(x_k)$, and $\lim_{i\to\infty} x_{k_i} = \lim_{i\to\infty} x_{k_i+1} = \bar{x}$, then $\bar{x}$ is optimal.

In order to apply Corollary 6 it is necessary to explore conditions under which $\lim_{i\to\infty} x_{k_i} = \lim_{i\to\infty} x_{k_i+1}$, and for this purpose the following notions are considered.

A real-valued function $f$ is *uniformly concave* [6] on a convex set $T$ if there exists a nondecreasing function $\delta(r) > 0$ on $(0,\infty)$ such that

$$f(\tfrac{1}{2}(x+y)) \geq \tfrac{1}{2}f(x) + \tfrac{1}{2}f(y) + \delta(|x-y|)$$

for all $x$, $y \in T$. A uniformly concave function is *strongly concave* [6] if $\delta(r) = \gamma r^2$ for some $\gamma > 0$. It is easily seen that if $T$ is compact, $f$ has continuous second partial derivatives on $T$, and the matrix of second partial derivatives of $f$ is negative definite at all points of $T$, then $f$ is strongly concave on $T$.

<u>Lemma 7</u>:

If $T$ is convex, $f$ is uniformly concave on $T$, and $\lim_{i\to\infty} x_{k_i} = \bar{x}$ then $\lim_{i\to\infty} x_{k_i+1} = \bar{x}$.

<u>Proof</u>:

Since $x_{k_i}$ maximizes the concave function $f$ over the convex set $S_{k_i} \cap T$ and $x_{k_i+1} \in T_{k_i+1} \cap T \subseteq S_{k_i} \cap T$, $f\left(x_{k_i}\right) \geq f\left(\tfrac{1}{2}\left(x_{k_i} + x_{k_i+1}\right)\right)$. Thus

$$f\left(x_{k_i}\right) \geq f\left(\tfrac{1}{2}\left(x_{k_i} + x_{k_i+1}\right)\right) \geq \tfrac{1}{2}f\left(x_{k_i}\right) + \tfrac{1}{2}f\left(x_{k_i+1}\right) + \delta\left(\left|x_{k_i} - x_{k_i+1}\right|\right),$$

$$f\left(x_{k_i}\right) - f\left(x_{k_i+1}\right) \geq 2\delta\left(\left|x_{k_i} - x_{k_i+1}\right|\right), \text{ and so}$$

$$(8) \quad \infty > f\left(x_{k_1}\right) - \max_{x\in S} f(x) \geq f\left(x_{k_1}\right) - \lim_{k\to\infty} f\left(x_{k_i}\right) = \sum_{i=1}^{\infty}\left(f\left(x_{k_i}\right) - f\left(x_{k_{i+1}}\right)\right)$$

$$\geq \sum_{i=1}^{\infty}\left(f\left(x_{k_i}\right) - f\left(x_{k_i+1}\right)\right) \geq 2\sum_{i=1}^{\infty}\delta\left(\left|x_{k_i} - x_{k_i+1}\right|\right)$$

But (8) implies that $\lim_{i\to\infty} \delta\left(\left|x_{k_i} - x_{k_i+1}\right|\right) = 0$, and since $\delta(\cdot)$ is positive and

nondecreasing on $(0,\infty)$ this implies that $\lim_{i\to\infty} |x_{k_i} - x_{k_i+1}| = 0$ . //

Clearly the maximum of a uniformly concave function on a convex set is unique if it is attained. Thus, Theorem 10 (which together with the hypotheses of Theorem 8 and the earlier assumption that $f$ attains its maximum on $T$ implies that $\{x_k\}$ is bounded), Corollary 6, and Lemma 7 yield the following result which allows inactive constraints to be dropped after each subproblem.

Theorem 8:

If $H(x)$ is determined by a limiting cutting-plane function, $H_k = H(x_k)$ , $T$ is convex, and $f$ is uniformly concave on $T$ , then $\{x_k\}$ converges to the unique maximum of $f$ on $S$ .

## 2. CONVERGENCE RATES

Levitin and Polyak [6] have established an arithmetic convergence rate for a cutting-plane algorithm which, when specialized to subsets of $E^n$, has $S_k = T_k$ (although their proof still holds if inactive constraints were dropped after each subproblem) and uses the cutting-plane method of Lemma 3. Here their algorithm and method of proof are generalized to show the same convergence rate for algorithms which allow inactive constraints to be dropped after each subproblem and for which the cutting plane at $x \in T - S$ may be generated at some point $w(x) \in T - S$ other than $x$ on the line segment joining $x$ to a point $t$ with $G(t) > 0$ (as in Lemma 4, although here, unfortunately, $w(x)$ must be bounded away from $S$) .

Theorem 9:

Suppose that $S = \{x : G(x) \geq 0\}$, $T$ is convex, $G(x)$ is concave on $T$, there exists $t$ with $G(t) > 0$, and for $x \in T - S$ define $\lambda(x) = \sup \{\lambda : \lambda x + (1 - \lambda)t \in S\}$ and set $w(x) = \alpha(x)x + (1 - \alpha(x))t$ for any $\alpha(x) \in [\lambda(x),1]$ with $G(w(x)) \leq \varepsilon G(x)$ where $0 < \varepsilon \leq 1$. Suppose also that for $P = \{w :$ there exists $x \in T - S$ with $w(x) = w\}$ there exists a function $\mu(w)$ from $P$ into $E^n$ with $|\mu(w)| \leq K$ for all $w \in P$ and such that $0 \leq G(w) + \mu(w)\cdot(y - w)$ for all $w \in P$ and $y \in S$,† and let $H_k = \{x : 0 \leq G(w(x_k)) + \mu(w(x_k)) \cdot (x - w(x_k))\}$. If $T$ is compact, $f$ is strongly concave and differentiable on $T$, and $\bar{x}$ is the unique maximum of $f$ on $S$, then for $k \geq 1$

$$f(x_k) - f(\bar{x}) \leq \frac{1}{a_1 k}$$

†See the remarks following Lemma 4.

and

$$|x_k - \bar{x}| \leq \frac{1}{a_2\sqrt{k}}$$

where

$$a_1 = 2\gamma\left(\frac{\varepsilon G(t)}{Kbd}\right)^2 \quad , \quad a_2 = \frac{2\gamma\varepsilon G(t)}{Kbd} \; ,$$

$d = \max \{|\nabla f(y)| \; : \; y \in T\}$ , $b = \max \{|y - t| \; : \; y \in T\}$ , and $\gamma$ is as in the definition of strong concavity.

Proof:

Let $\lambda_k = \lambda(x_k)$ , $\alpha_k = \alpha(x_k)$ , $w_k = w(x_k)$ , and $\mu_k = \mu(w_k)$ . Clearly, $G(\lambda_k x_k + (1 - \lambda_k)t) = 0$ and so by concavity

$$0 = G(\lambda_k x_k + (1 - \lambda_k)t) \geq \lambda_k G(x_k) + (1 - \lambda_k)G(t)$$

and

$$-G(x_k) \geq \frac{1}{\lambda_k}(1 - \lambda_k)G(t) \geq (1 - \lambda_k)G(t) \; . \tag{9}$$

Since $t \in S$ ,

$$0 \leq G(w_k) + \mu_k \cdot (t - w_k) \; . \tag{10}$$

But $x_k$ is the unique maximum of $f$ on $S_k \cap T$ and it is easily seen that $x_k \notin H_k$ so $x_k \neq x_{k+1} \in T_{k+1} \cap T = H_k \cap S_k \cap T$ and by the strict concavity of $f$ on $T$ ,

$$0 = G(w_k) + \mu_k \cdot (x_{k+1} - w_k) \tag{11}$$

or

(12) $$0 = G(w_k) + \alpha_k\mu_k\cdot(x_{k+1} - x_k) + (1 - \alpha_k)\mu_k\cdot(x_{k+1} - t) .$$

By (10) and (11)

(13) $$0 \geq \mu_k\cdot(x_{k+1} - t) ,$$

so by (12) and (13)

(14) $$0 \leq G(w_k) + \alpha_k\mu_k\cdot(x_{k+1} - x_k) \leq \varepsilon G(x_k) + K|x_{k+1} - x_k| .$$

By the concavity of $f$ and the optimality of $\bar{x}$,

(15) $$\begin{aligned} f(x_k) - f(\bar{x}) &\leq f(x_k) - f(\lambda_k x_k + (1 - \lambda_k)t) \\ &\leq (1 - \lambda_k)(x_k - t)\cdot\nabla f(\lambda_k x_k + (1 - \lambda_k)t) \\ &\leq (1 - \lambda_k)|x_k - t|\cdot|\nabla f(\lambda_k x_k + (1 - \lambda_k)t)| \\ &\leq (1 - \lambda_k)bd . \end{aligned}$$

Combining (15), (9), and (14),

(16) $$\begin{aligned} f(x_k) - f(\bar{x}) &\leq (1 - \lambda_k)bd \\ &\leq \frac{1}{G(t)}(-G(x_k))bd \\ &\leq \frac{Kbd}{\varepsilon G(t)}|x_k - x_{k+1}| . \end{aligned}$$

Since $x_k$, $x_{k+1} \in S_k \cap T$ and $x_k$ maximizes the strongly concave function $f$ over the convex set $S_k \cap T$, for some $\gamma > 0$

(17) $$f(x_k) > f(\tfrac{1}{2}(x_k + x_{k+1})) \geq \tfrac{1}{2}f(x_k) + \tfrac{1}{2}f(x_{k+1}) + \gamma|x_k - x_{k+1}|^2$$

and from (17)

(18) $$f(x_k) - f(x_{k+1}) \geq 2\gamma|x_k - x_{k+1}|^2 .$$

Now let $D_k = f(x_k) - f(\bar{x}) > 0$ . By (16) and (18)

$$D_k^2 \leq \left(\frac{Kbd}{\varepsilon G(t)}\right)^2 |x_k - x_{k+1}|^2 \leq \left(\frac{Kbd}{\varepsilon G(t)}\right)^2 \left(\frac{1}{2\gamma}\right)(D_k - D_{k+1})$$

or

(19) $$D_{k+1} \leq D_k - a_1 D_k^2 = D_k(1 - a_1 D_k) .$$

The arithmetic convergence rate for $D_k$ then follows from (19) as in [3] by observing that

(20) $$\frac{1}{D_{k+1}} \geq \frac{1}{D_k}\left(\frac{1}{1 - a_1 D_k}\right) = \frac{1}{D_k}\left(\sum_{i=0}^{\infty} (a_1 D_k)^i\right)$$

$$\geq \frac{1}{D_k}(1 + a_1 D_k) = \frac{1}{D_k} + a_1$$

and using induction on (20) to get

$$\frac{1}{D_k} \geq \frac{1}{D_o} + a_1 k$$

or

(21) $$D_k \leq \frac{1}{\frac{1}{D_o} + a_1 k} \leq \frac{1}{a_1 k} .$$

As in (17) and (18), it follows that

(22) $$D_k = f(x_k) - f(\bar{x}) \geq 2\gamma|x_k - \bar{x}|^2 ,$$

and by (21) and (22)

$$|x_k - \bar{x}| \leq \frac{1}{\sqrt{2\gamma a_1 k}} \ . \ ||$$

## 3. SOME USEFUL ANALYTIC RESULTS

In the introduction to this paper, the somewhat unintuitive assumption was made that one is given a function $f$ on a closed set $T$ such that $f$ attains its maximum on any nonempty closed subset of $T$. Corollary 11 provides intuitively appealing conditions for this to be true. Theorem 10 was proven by Zoutendijk [11] under the additional assumptions that $T = E^n$ and $f$ is differentiable on $T$.

**Theorem 10:**

If $T$ is a closed convex subset of $E^n$, $f$ is concave and upper semi-continuous on $T$, and the set of maxima of $f$ on $T$ is nonempty and bounded, then $\{x : x \in T, f(x) \geq \alpha\}$ is bounded for all $\alpha$.

**Proof:**

Let $W$ be the set of maxima of $f$ on $T$. Since $f$ is upper semi-continuous on the closed set $T$, it follows that $W$ is closed and thus $W$ is compact and nonempty by hypothesis. Now pick any $\bar{x} \in W$ and any $\gamma > 0$ such that $x \in W$ implies $|x - \bar{x}| < \gamma$. Let $B = \{x : x \in T, |x - \bar{x}| = \gamma\}$. If $B = \emptyset$, then $T$ is bounded by convexity and the proof is complete, so assume $B \neq \emptyset$. Let $M = \sup_{x \in B} f(x)$. Since $f$ is upper semi-continuous on the nonempty compact set $B$ it follows that $f$ attains its maximum on $B$, so $f(\bar{x}) > M$ since $W \cap B = \emptyset$. Now pick any $\alpha < f(\bar{x})$ (the result follows trivially for $\alpha \geq f(\bar{x})$) and any $x \in T$ with $|x - \bar{x}| > \gamma$ and $f(x) \geq \alpha$. Let $\lambda = \dfrac{\gamma}{|x - \bar{x}|} \in (0,1)$ and $z = \bar{x} + \lambda(x - \bar{x})$. Then $z \in T$ by convexity and $|z - \bar{x}| = \lambda|x - \bar{x}| = \gamma$ so $z \in B$. Thus by concavity

$$M \geq f(z) \geq \lambda f(x) + (1 - \lambda)f(\bar{x}) \geq \lambda\alpha + (1 - \lambda)f(\bar{x}) \tag{23}$$

and substituting $\lambda = \dfrac{\gamma}{|x - \bar{x}|}$ into (23),

$$\frac{\gamma(f(\bar{x}) - \alpha)}{f(\bar{x}) - M} \geq |x - \bar{x}|$$

and $\{x : x \in T , f(x) \geq \alpha\}$ must be bounded. ||

Corollary 11:

If $T$ is a closed convex subset of $E^n$ , $S$ is a nonempty closed subset of $T$ , $f$ is concave and upper semi-continuous on $T$ , and the set of maxima of $f$ on $T$ is nonempty and bounded, then $f$ attains its maximum on $S$ . Hence, a strictly concave upper semi-continuous function which attains its maximum on $E^n$ (such as any strictly concave quadratic function) will attain its maximum on any nonempty closed subset of $E^n$ .

Proof:

Pick any $y \in S$ . Then it suffices to show that $S_o = \{x : x \in S , f(x) \geq f(y)\}$ is bounded. But $S_o \subseteq \{x : x \in T , f(x) \geq f(y)\}$ which is bounded by Theorem 10. ||

In the introduction, it was mentioned that if each $S_k$ is the intersection of $E^n$ and all the linear constraints of $T_k$ that are active at $x_k$ then each $T_k$ is the intersection of no more than n+1 closed half-spaces. This follows directly by applying the following result and induction.

Lemma 12:

Suppose $a_i \in E^n$ and $b_i \in E^1$ for $i = 1, \ldots, m+1$ , and $x$ , $z \in E^n$ . If $a_1, \ldots, a_m$ are independent (so $m \leq n$) ,

(24) $$a_i \cdot x = b_i \qquad \text{for } i = 1, \ldots, m ,$$

(25) $$a_{m+1} \cdot x < b_{m+1} ,$$

and $I = \{i : a_i \cdot z = b_i , 1 \le i \le m + 1\}$ , then $\{a_i : i \in I\}$ is independent (and hence has no more than $n$ elements).

Proof:

Suppose $\{a_i : i \in I\}$ is dependent. Then there exist numbers $\lambda_i$ , $i \in I$ , not all $0$ , with $\sum_{i\in I} \lambda_i a_i = 0$ . Since $a_1, \ldots, a_m$ are independent, $m + 1 \in I$ and $\lambda_{m+1} \neq 0$ and so we may suppose $\lambda_{m+1} > 0$ . Then

$$(26) \qquad 0 = \sum_{i\in I} (\lambda_i a_i)\cdot z = \sum_{i\in I} \lambda_i (a_i \cdot z) = \sum_{i\in I} \lambda_i b_i .$$

But by (24), (25), and $\lambda_{m+1} > 0$ ,

$$0 = \sum_{i\in I} (\lambda_i a_i)\cdot x = \sum_{i\in I-\{m+1\}} \lambda_i b_i + \lambda_{m+1} a_{m+1}\cdot x < \sum_{i\in I} \lambda_i b_i$$

which contradicts (26). ||

## 4. RELATED WORK

Van Slyke [8] has considered the result of Theorem 8 in spaces more general than $E^n$ and has applied it to an optimal control problem.

C. Eaves and W. Zangwill have recently informed me that they are currently developing a theory for cutting-plane algorithms which allows inactive constraints to be dropped after each subproblem. There seems to be some overlap between their results and mine, although our approaches are different.

## REFERENCES

[1] Dantzig, G. B., LINEAR PROGRAMMING AND EXTENSIONS, Princeton University Press, (1963).

[2] Dantzig, G. B. and R. W. Cottle, "Positive (Semi-) Definite Programming," in NONLINEAR PROGRAMMING (J. Abadie, ed.), John Wiley & Sons, Inc., pp. 55-73, (1967).

[3] Frank, M. and P. Wolfe, "An Algorithm for Quadratic Programming," Naval Research Logistics Quarterly, Vol. 3, pp. 95-110, (1956).

[4] Kelley, J. E., Jr., "The Cutting-Plane Method for Solving Convex Programs," Journal of the Society for Industrial and Applied Mathematics, Vol. 8, pp. 703-712, (1960).

[5] Lemke, C. E., "Bimatrix Equilibrium Points and Mathematical Programming," Management Science, Vol. 11, pp. 442-453, (1955).

[6] Levitin, E. S. and B. T. Polyak, "Constrained Minimization Methods," Zh. vychisl. Mat. mat. Fiz. (in Russian), Vol. 6, pp. 787-823, (1966); also, U.S.S.R. Computational Mathematics and Mathematical Physics (in English), pp. 1-50, (1968).

[7] Mangasarian, O. L., NONLINEAR PROGRAMMING, to appear, (1968).

[8] Van Slyke, R., "Cutting-Plane Algorithms and State Space Constrained Linear Optimal Control Problems," to appear, (1969).

[9] Veinott, A. F., Jr., "The Supporting Hyperplane Method for Unimodal Programming," Operations Research, Vol. 15, pp. 147-152, (1967).

[10] Zangwill, W. J., NONLINEAR PROGRAMMING: A UNIFIED APPROACH, Prentice Hall, Inc., (1969).

[11] Zoutendijk, G., METHODS OF FEASIBLE DIRECTIONS, Elsevier Publishing Company, (1960).

[12] Zoutendijk, G., "Nonlinear Programming: A Numerical Survey," Journal on Control of the Society for Industrial and Applied Mathematics, Vol. 4, pp. 194-210, (1966).

Unclassified
Security Classification

# DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

| | |
|---|---|
| 1. ORIGINATING ACTIVITY (Corporate author) University of California, Berkeley | 2a. REPORT SECURITY CLASSIFICATION Unclassified |
| | 2b. GROUP |
| 3. REPORT TITLE CUTTING-PLANE METHODS WITHOUT NESTED CONSTRAINT SETS | |
| 4. DESCRIPTIVE NOTES (Type of report and inclusive dates) Research Report | |
| 5. AUTHOR(S) (First name, middle initial, last name) Donald M. Topkis | |
| 6. REPORT DATE June 1969 | 7a. TOTAL NO. OF PAGES 19 — 7b. NO. OF REFS 12 |
| 8a. CONTRACT OR GRANT NO. Nonr-222(83) | 9a. ORIGINATOR'S REPORT NUMBER(S) ORC 69-14 |
| b. PROJECT NO. NR 047 033 | |
| c. Research Project No.: RR 003 07 01 | 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) |
| d. | |
| 10. DISTRIBUTION STATEMENT This document has been approved for public release and sale; its distribution is unlimited. | |
| 11. SUPPLEMENTARY NOTES Also supported by the National Science Foundation under Grant GP-8695. | 12. SPONSORING MILITARY ACTIVITY MATHEMATICAL SCIENCES DIVISION |

13. ABSTRACT

SEE ABSTRACT.

DD FORM 1 NOV 65 1473 (PAGE 1)

S/N 0101-807-6811

Unclassified
Security Classification

A-31408

Unclassified
Security Classification

| 14 KEY WORDS | LINK A | | LINK B | | LINK C | |
|---|---|---|---|---|---|---|
| | ROLE | WT | ROLE | WT | ROLE | WT |
| Nonlinear Programming | | | | | | |
| Algorithms | | | | | | |
| Cutting-Plane Algorithms | | | | | | |
| Convergence Rates | | | | | | |
| Uniformly and Strongly Concave Functions | | | | | | |